

# Nachvollziehbarkeit mit Hinblick auf Privacy-Verletzungen

**Andreas Schreiber**

Intelligente und Verteilte Systeme

DLR Köln/Braunschweig/Berlin



Wissen für Morgen



# Fragen

## Viele Informationen stecken in den Metadaten

- Wo werden Daten gespeichert?
- Wie werden Daten verarbeitet?
- Wer hat Zugriff auf die Daten?

## Metadaten des Datenverarbeitungsprozesses: Provenance

- **Nachvollziehbarkeit** und **Reproduzierbarkeit**
- Standardformat PROV des W3C für **Interoperabilität**



# Nachvollziehbarkeit von Prozessen

Provenance is

**information about entities, activities, and people**  
involved in  
**producing a piece of data or thing,**  
which can be used to form  
**assessments about its quality, reliability or trustworthiness.**

*PROV W3C Working Group*

<https://www.w3.org/TR/prov-overview>



# Kern-Elemente von Provenance

## Entities

- Physikalische Objekte, digitale Objekte, Konzepte oder sonstiges
- Dateien, Dokumente, Webseiten, Grafiken



## Activities

- Erzeugen oder verarbeiten Entities
- Einzelne Aktionen, Operationen oder Prozesse



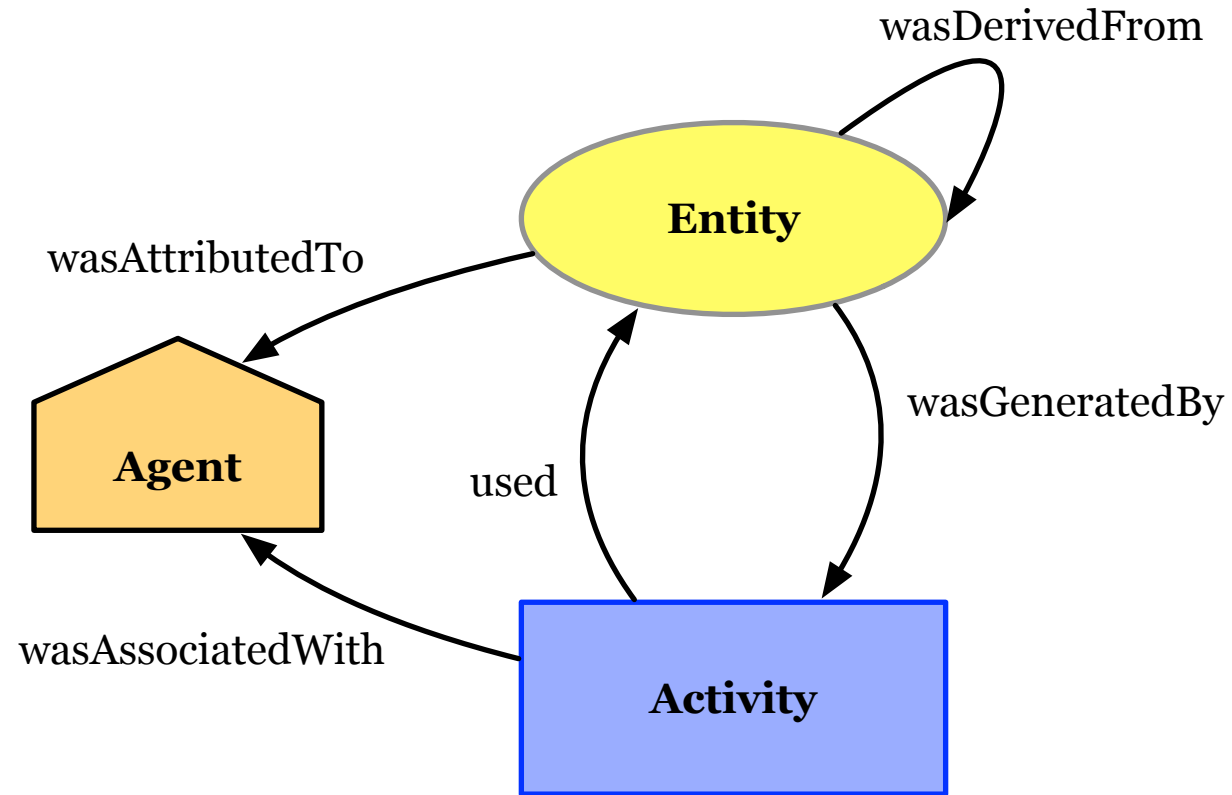
## Agents

- Agents haben bestimmte Rollen in Bezug auf die Activities; sie sind im gewissen Sinne für Activities verantwortlich
- Menschen, Software oder Organisationen

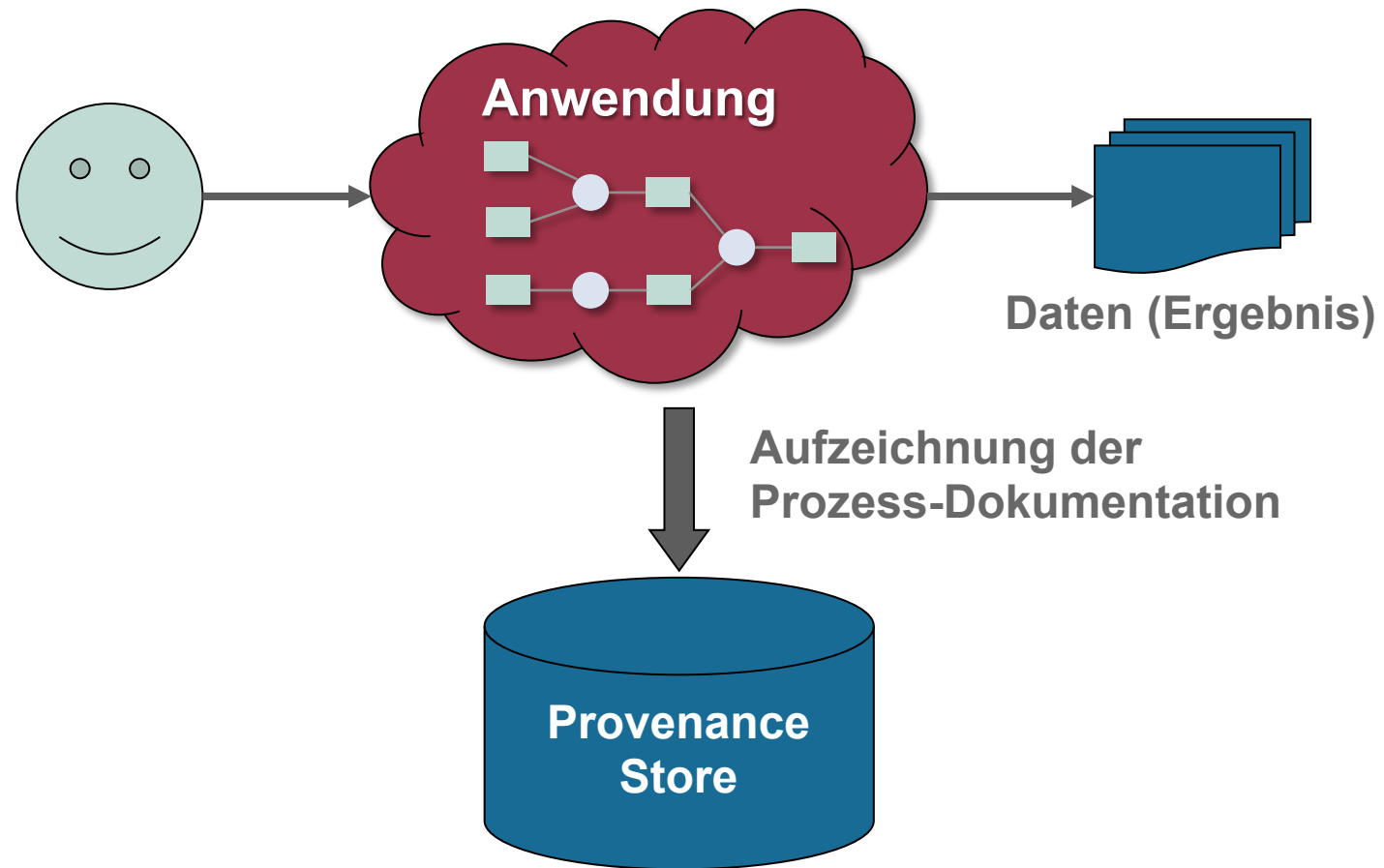




# Provenance-Modellierung

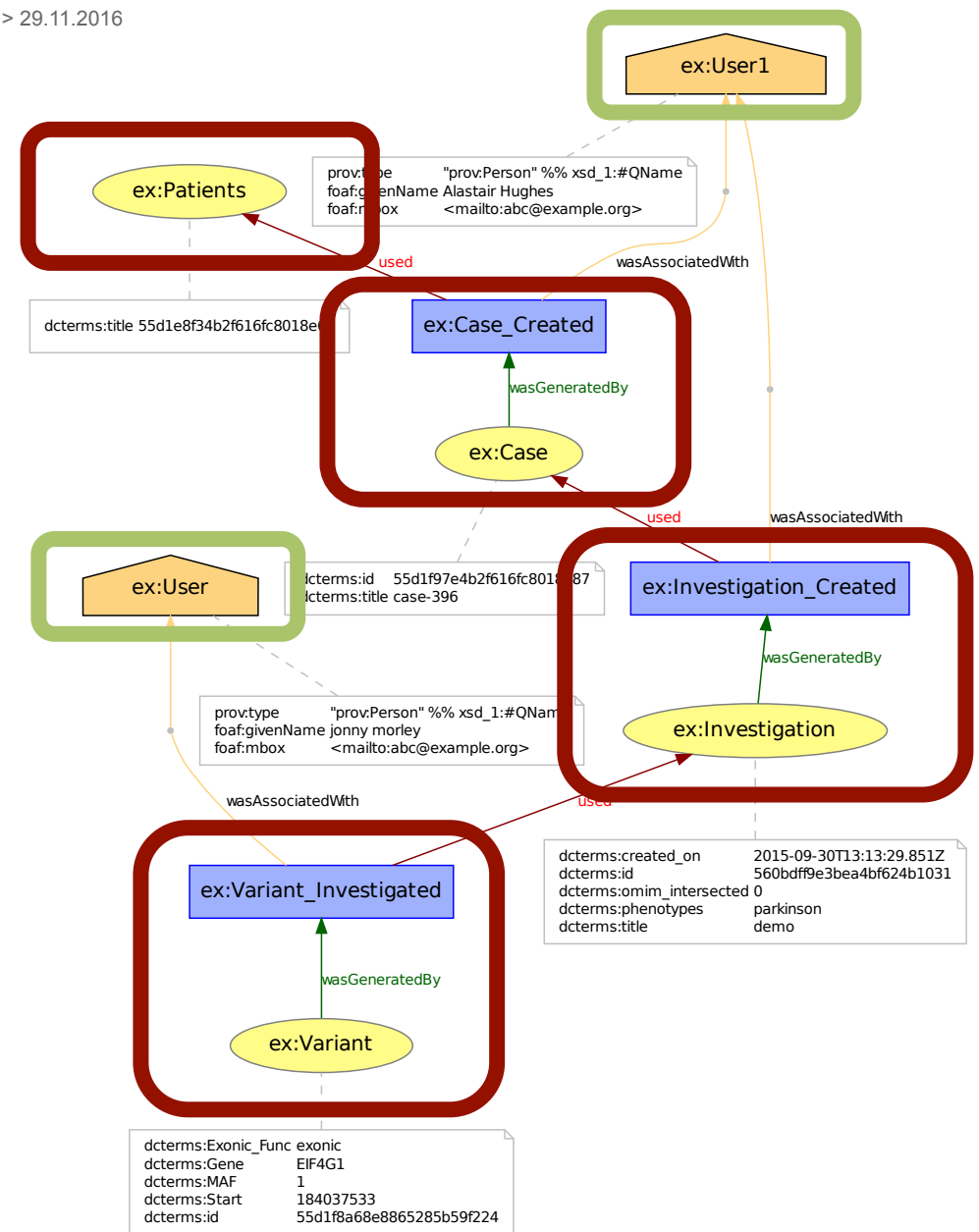
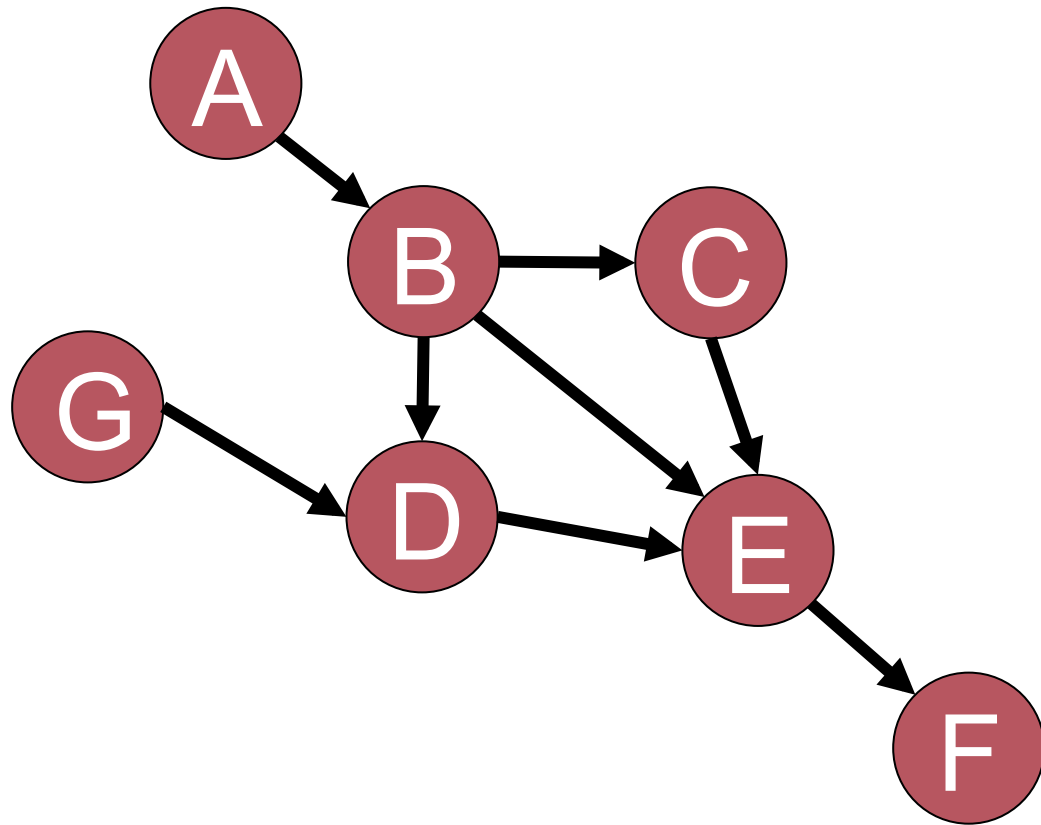


# Speichern der Provenance-Informationen

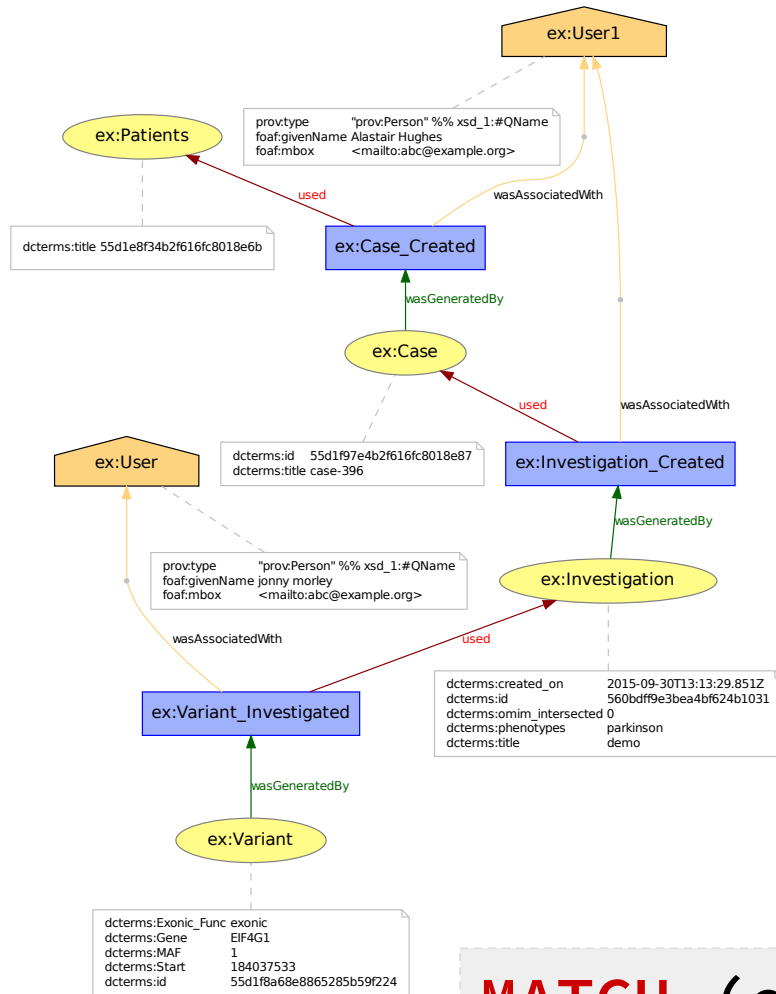


# Graphen

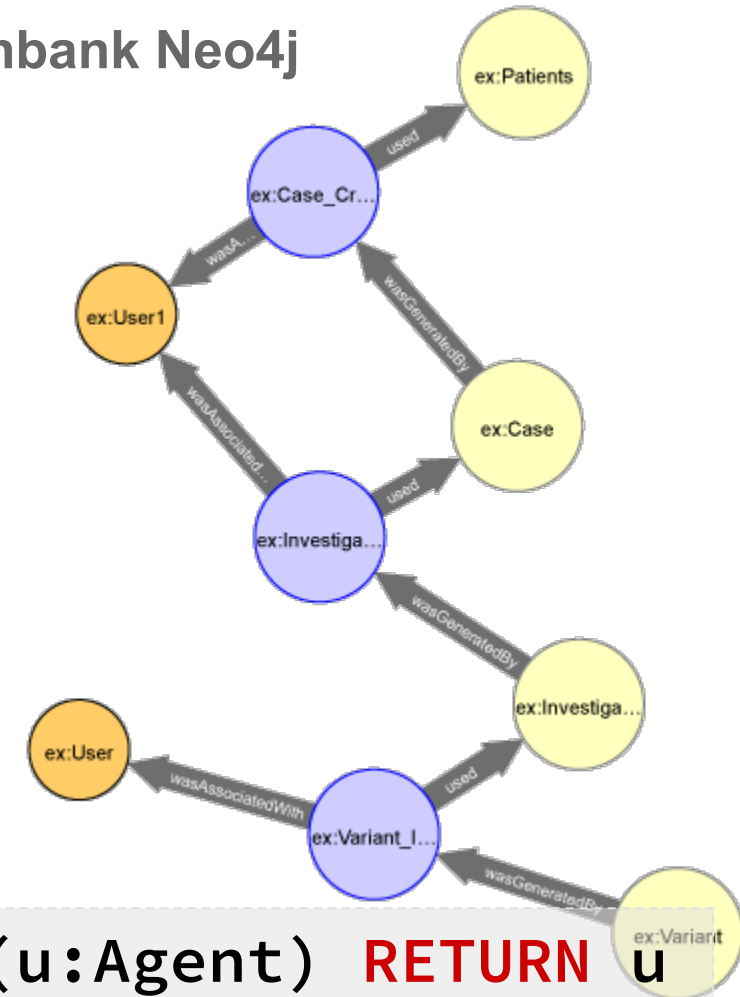
Provenance ist ein Directed Acyclic Graph (DAG)



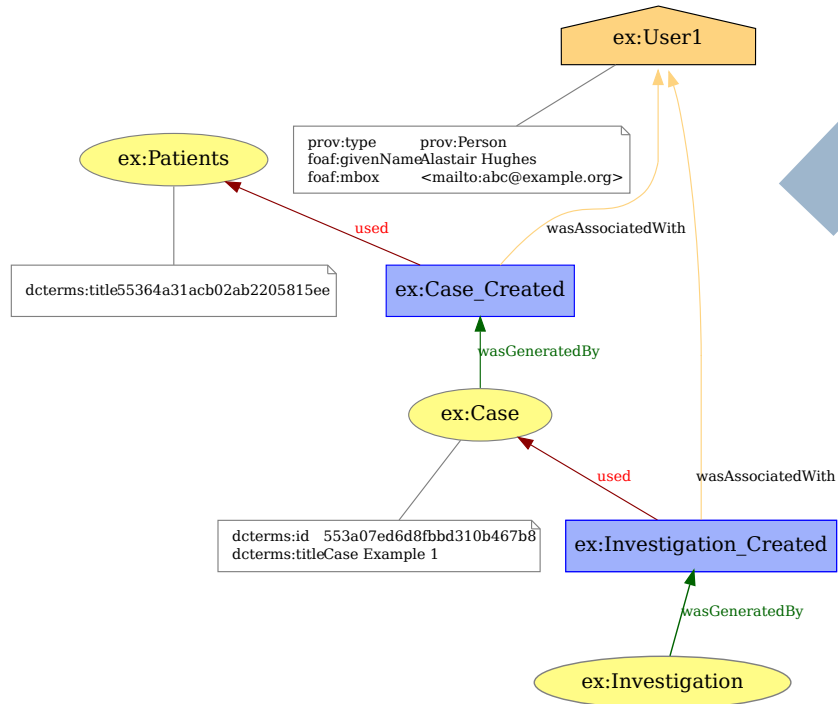
# Graph-Datenbanken



## Graph-Datenbank Neo4j

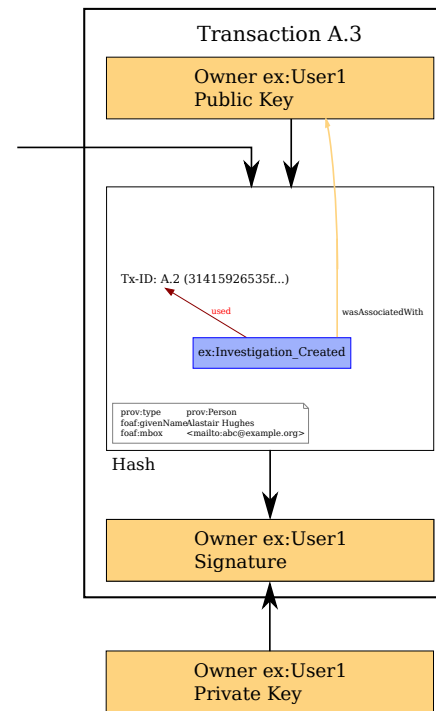


# Blockchain-enabled Data Science

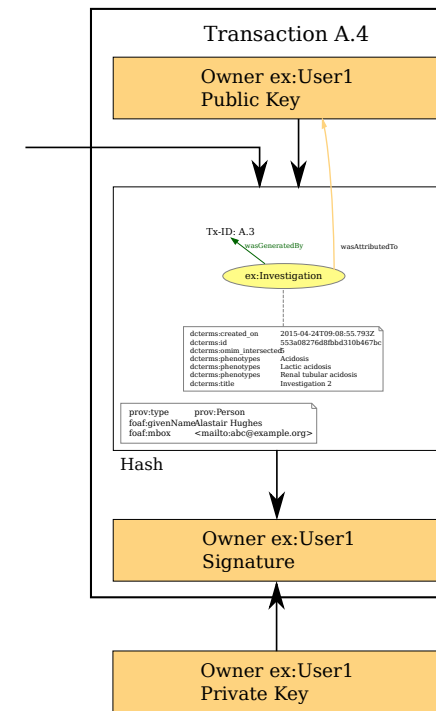


Document A

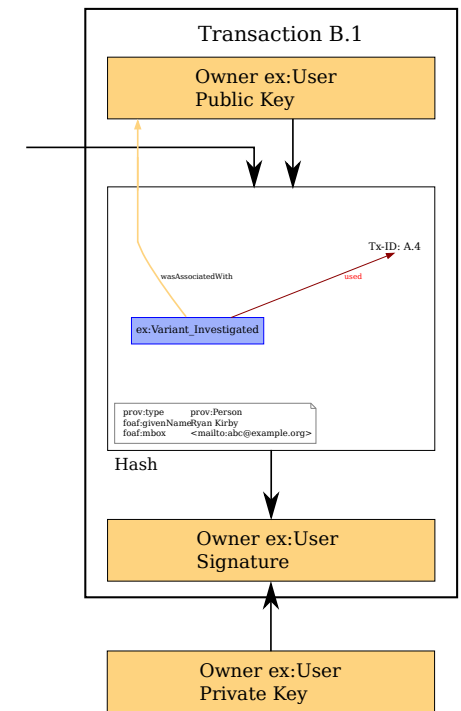
## Blockchain: Dezentrale und unveränderliche „Big Data“-Datenbank



Create Asset



Create Asset



Create Asset



# Anwendungskontext: Medizin und Gesundheitsdaten



Medizin- und Gesundheits-Apps



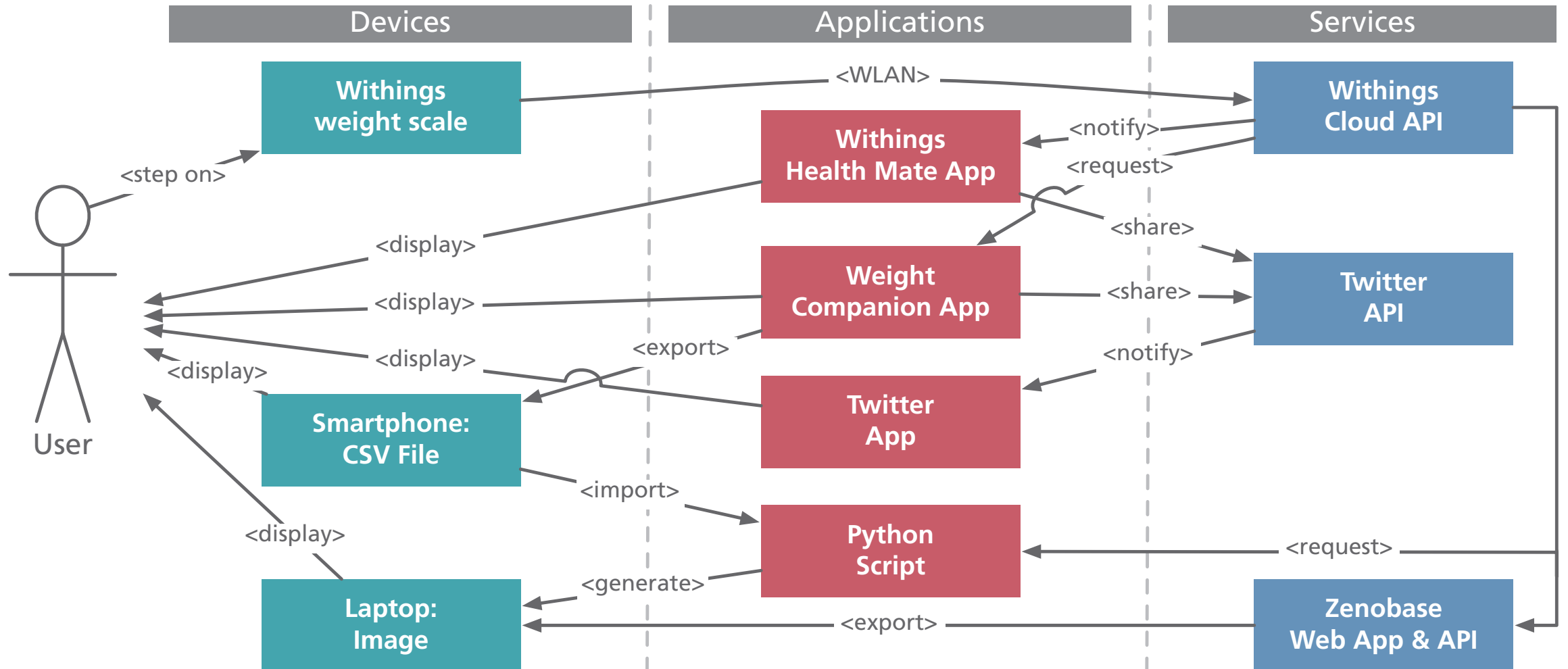
Medizinische Studien



Telemedizin und AAL



# Beispiel-Datenworkflow



# Provenance

```
Date,Time,Weight,Waist,Hip,Device,Comment
"Jun 13, 2012",14:00,83.7,,Withings,
"Jun 13, 2012",14:08,79.7,,Withings,
"Jun 15, 2012",21:59,82.7,,Withings,
"Jun 15, 2012",22:04,82.7,,Withings,
"Jun 24, 2012",18:32,86.1,,Withings,
"Jun 26, 2012",07:42,80.8,,Withings,
"Jun 27, 2012",07:40,81.1,,Withings,
"Jun 29, 2012",07:34,79.4,,Withings,
"Jun 30, 2012",22:12,81.7,,Withings,
"Jul 1, 2012",11:21,80.6,,Withings,
"Jul 7, 2012",17:04,80.7,,Withings,
"Jul 10, 2012",07:46,81.8,,Withings,
"Jul 11, 2012",07:32,78.6,,Withings,
"Jul 12, 2012",07:26,79.4,,Withings,
```

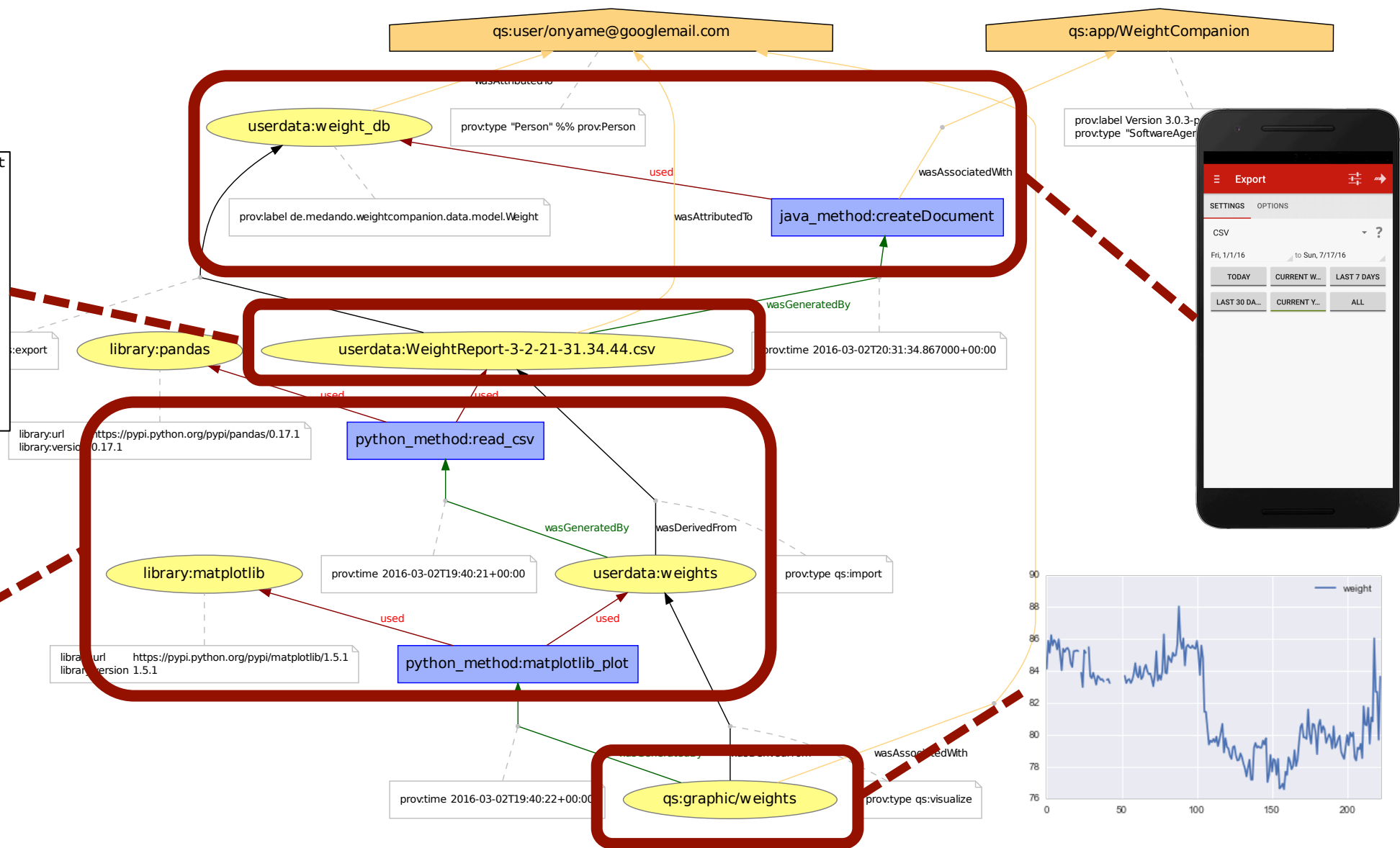
```
# Application Import
from pandas import DataFrame, Series, read_csv
import matplotlib.pyplot as plt

prov.entity('library:pandas', {'library:version': '0.17.1'})
prov.entity('library:matplotlib', {'library:version': '1.5.1'})

# Import weights from CSV file
WC_FILE = 'WeightReport-3-2-21-31.34.44.csv'
entity_csvfile = prov.entity('userdata:weights', {'prov:label': 'WeightReport-3-2-21-31.34.44.csv'})
prov.wasAttributedTo(entity_csvfile, agent_us)
wc_data = read_csv(WC_FILE, parse_dates={'Timestamp': [0, 1]}, index_col='Timestamp')

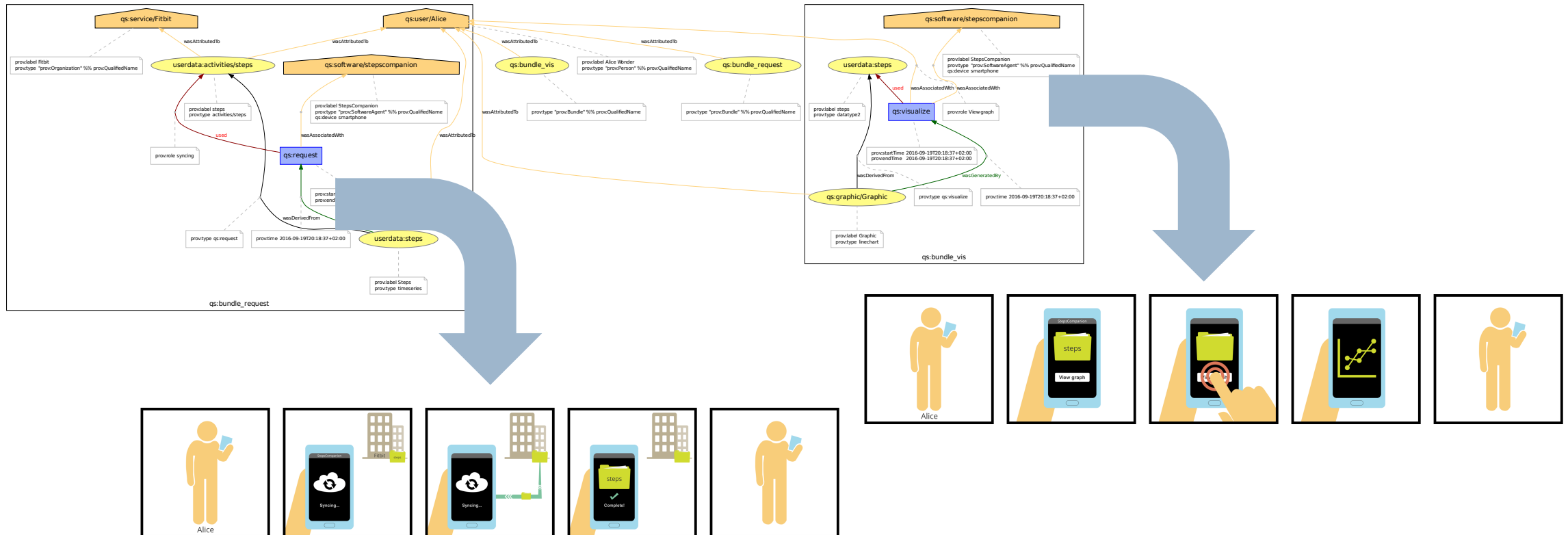
# Get just the weights
weights = wc_data[['Weight']]
entity_weights = prov.entity('userdata:weights', {'prov:label': 'WeightReport-3-2-21-31.34.44.csv'})
prov.activity('python_method:read_csv', {'prov:time': '2016-03-02T19:40:21+00:00'})
prov.wasGeneratedBy(entity_weights, 'python_method:read_csv')
prov.used('python_method:read_csv', entity_csvfile)
prov.wasDerivedFrom(entity_weights, entity_csvfile)
other_attributes={'prov:type': 'qs:import'}
prov.used('python_method:read_csv', 'library:pandas')

# Plot the weights
weights.plot(title = 'Weight', legend=False)
entity_plot_weights = prov.entity('graphic:weights', {'prov:time': '2016-03-02T19:40:22+00:00'})
prov.wasGeneratedBy(entity_plot_weights, 'python_method:matplotlib_plot')
prov.used('python_method:matplotlib_plot', entity_weights)
prov.wasAssociatedWith(entity_plot_weights, 'qs:graphic/weights')
prov.type('qs:visualize')
```



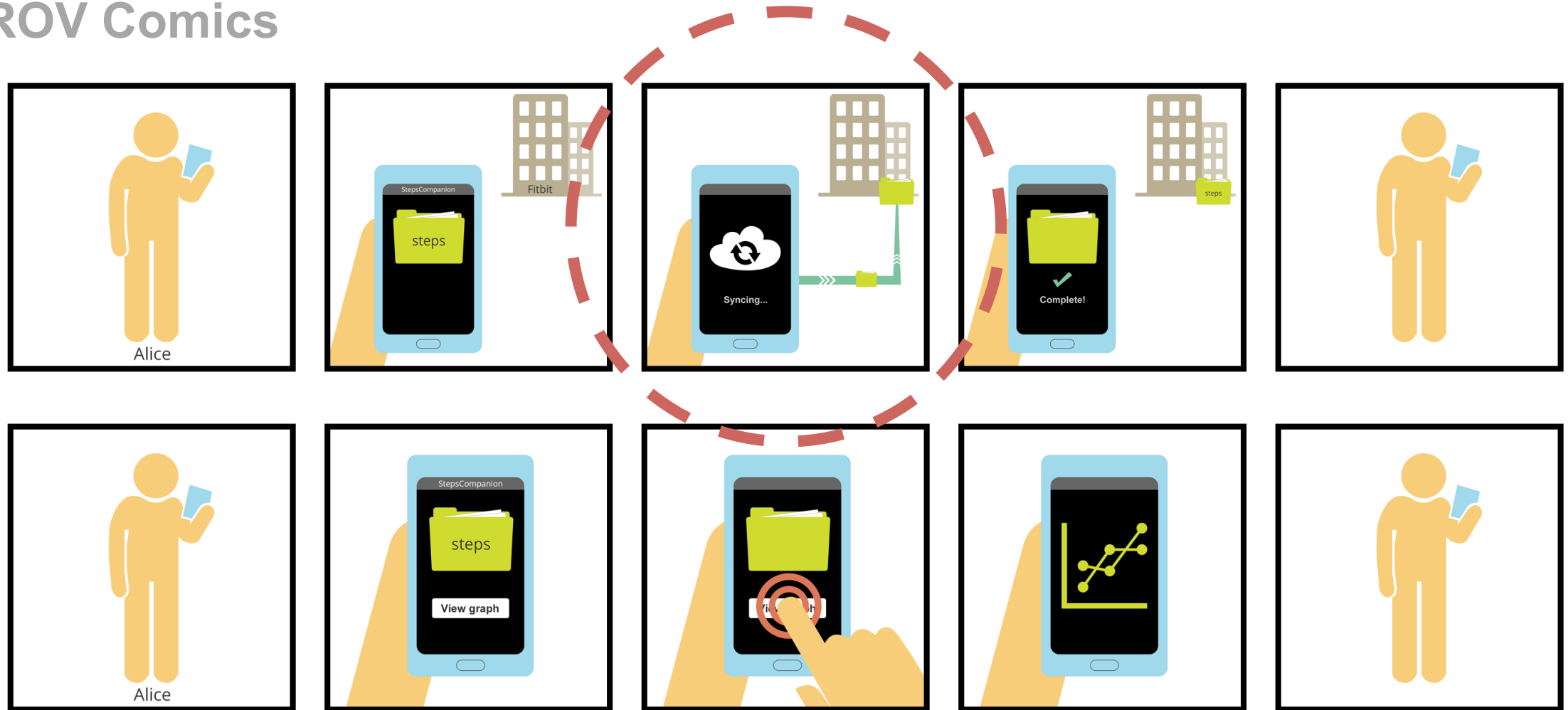
# Provenance for Humans

## PROV Comics



# Provenance for Humans

## PROV Comics





# Vielen Dank!

Fragen?

[Andreas.Schreiber@dlr.de](mailto:Andreas.Schreiber@dlr.de)  
[www.DLR.de/sc](http://www.DLR.de/sc) | [@onyame](#)